PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : H04L 9/00

A1

(11) International Publication Number: WO 92/03000

(43) International Publication Date: 20 February 1992 (20.02.92)

(21) International Application Number: PCT/US91/05386

(22) International Filing Date: 30 July 1991 (30.07.91)

(30) Priority data:

561,888	2 August 1990 (02.08.90)	US
666,896	8 March 1991 (08.03.91)	US

(71) Applicant: BELL COMMUNICATIONS RESEARCH, INC. [US/US]; 290 West Mount Pleasant Avenue, Livingston, NJ 07039-2729 (US).

(72) Inventors: HABER, Stuart, Alan ; 22 Irving Place, Apt. 2C, New York, NY 10003 (US). STORNETTA, Wakefield, Scott, Jr. ; 34 Harding Terrace, Morristown, NJ 07960 (US).

(74) Agents: WINTER, Richard, C.; PCT Int'l., Inc., c/o Bell Communications Research, Inc., International Coordinator, Room 2E-304, 290 West Mount Pleasant Avenue, Livingston, NJ 07039 (US) et al.

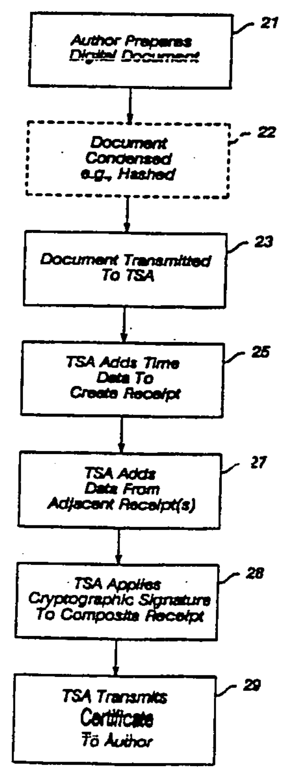
(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: METHOD FOR SECURE TIME-STAMPING OF DIGITAL DOCUMENTS

(57) Abstract

A system for time-stamping a digital document is disclosed which protects the secrecy of the document text and provides a tamper-proof time seal establishing an author's claim to the temporal existence of the document. Initially the author prepares the document (21), which may then be condensed by a process such as hashing (22). Next, the document is transmitted to the Time Stamping Authority (23), which adds time data to create a receipt (25) and data from adjacent receipts (27). Thereafter, the Time Stamping Authority applies a cryptographic signature to the composite receipt (28), which is then transmitted to the author (29).



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU+	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

\+ It is not yet known for which States of the former Soviet Union any designation of the Soviet Union has effect.

METHOD FOR SECURE TIME-STAMPING OF DIGITAL DOCUMENTS

BACKGROUND OF THE INVENTION

In many situations there is a need to establish the date on which a document was created and to prove that the text of a document in question is in fact the same as that of the original dated document. For example, in intellectual property matters it is often crucial to verify the date on which a person first put into writing the substance of an invention. A common procedure for thus "time-stamping" an inventive concept comprises daily notations of one's work in a laboratory notebook. Indelibly dated and signed entries are made one after another on each page of the notebook where the sequentially numbered, sewn-in pages make it difficult to revise the record without leaving telltale signs. The validity of the record is further enhanced by the regular review and signed witnessing by a generally disinterested third party. Should the time of the concept become a matter for later proof, both the physical substance of the notebook and the established recording procedure serve as effective evidence in substantiating the fact that the concept existed at least as early as the notebook witness date.

The increasingly widespread use of electronic documents, which include not only digital representations of readable text but also of video, audio, and pictorial data, now poses a serious threat to the viability of the "notebook" concept of establishing the date of any such document. Because electronic digital documents are so

easily revised, and since such revisions may be made without telltale sign, there is available limited credible evidence that a given document truly states the date on which it was created or the message it originally carried. For the same reasons there even arises serious doubt as to the authenticity of a verifying signature. Without an effective procedure for ensuring against the surreptitious revision of digital documents, a basic lack of system credibility prevents the efficiencies of electronic documentation from being more widely implemented.

Some procedures are presently available for verifying electronic document transmissions; however, such procedures are limited in application to bilateral communications. That is, in such communications the sender essentially desires to verify to the receiver the source and original content of the transmitted document. For example, "private key" cryptographic schemes have long been employed for message transmission between or among a limited universe of individuals who are known to one another and who alone know the decrypting key. Encryption of the message ensures against tampering, and the fact that application of the private key reveals the "plaintext" of the transmitted message serves as proof that the message was transmitted by one of the defined universe. The time of creation of the message is only collaterally established, however, as being not later than its receipt by the addressee. This practice thus fails to provide time-stamp evidence that would be useful in an unlimited universe at a later date.

A more broadly applicable verifying communication procedure, that of "public key" cryptography, has been described by Diffie and Hellman ("New Directions in Cryptography", IEEE Transactions On Information Theory,

Vol. IT-22, November 1976, pp. 644-654) and more recently implemented by Rivest et al. in U.S. Patent 4,405,829, issued 20 September 1983. While this scheme expands the utilizing universe to a substantially unlimited number of system subscribers who are unknown to one another, but for a public directory, verifiable communications remain bilateral. These limitations persist, since although a public key "signature", such as that which entails public key decryption of a message encrypted with the private key of the transmitter, provides any member of the unlimited universe with significant evidence of the identity of the transmitter of the message, only a given message recipient can be satisfied that the message existed at least as early as the time of its receipt. Such receipt does not, however, provide the whole universe with direct evidence of time of the message's existence. Testimony of a such a recipient in conjunction with the received message could advance the proof of message content and time of its existence, but such evidence falls victim to the basic problem of ready manipulation of electronic digital document content, whether by originator or witness.

Thus, the prospect of a world in which all documents are in easily modifiable digital form threatens the very substance of existing procedures for establishing the credibility of such documents. There is clearly a significant present need for a system of verification by which a digital document may be so fixed in time and content that it can present, at least to the extent currently recognized in tangible documents, direct evidence on those issues.

SUMMARY OF THE INVENTION

The present invention yields such a reliable system in a method of time-stamping digital documents that provides the equivalent of two essential characteristics of accepted document verification. First, the content of a document and a time stamp of its existence are "indelibly" incorporated into the digital data of the document so that it is not possible to change any bit of the resulting time-stamped data without such a change being apparent. In this manner, the state of the document text is fixed at the instant of time-stamping. Second, the time at which the digital document is stamped is verified by a "witnessing" digital signature procedure that deters the incorporation of a false time statement. In essence, the method transfers control of the time-stamping step from the author to an independent agent and removes from the author the ability to influence the agent in the application of other than a truthful time stamp.

The method of the present invention presumes a number of document authors distributed throughout a communication network. Such authors may be individuals, companies, company departments, etc. each representing a distinct and identifiable, e.g. by ID number or the like, member of the author universe. In one embodiment of the invention, this universe may constitute the clientele of a time-stamping agency (TSA), while in another embodiment the distributed authors may serve as agents individually performing the time-stamping service for other members of the universe.

In its general application, as depicted in FIG. 1 of the drawing, the present method entails an author's

preparation of a digital document, which may broadly comprise any alphanumeric, audio, or pictorial presentation, and the transmission of the document, preferably in a condensed representative form, to the TSA. The TSA time-stamps the document by adding digital data signifying the current time, applying the agency's cryptographic signature scheme to the document, and transmitting the resulting document, now a certificate of the temporal existence of the original document, back to the author where it is held for later use in required proof of such existence. Alternatively, the TSA may time-stamp the document to create a receipt by adding digital data signifying the current time, concatenate the receipt with the current cryptographic catenation of its prior time stamp receipts, and create a new catenation from the composite document by means of a deterministic function, such as discussed in greater detail below. The resulting catenate value is then included with time and other identifying data to yield the certificate.

To ensure against interception of confidential document information during transmission to the TSA, and to reduce the digital bandwidth required for transmission of an entire document, the author may optionally convert the digital document string to a unique value having vastly condensed digital size by means of a deterministic function which may, for example, be any one of a number of algorithms known in the art as "one-way hash functions". Such an application of hash functions has been described, among others, by Damgard in his discussions on the improvement of security in document signing techniques ("Collision-Free Hash Functions and Public Key Signature Schemes", Advances in Cryptology -- Eurocrypt '87, Springer-Verlag, LNCS, 1988, Vol. 304, pp. 203-217). In practice of the present invention, however, the "one-way"

characteristic typical of a hashing algorithm serves an additional purpose; that is, to provide assurance that the document cannot be secretly revised subsequent to the time the TSA applies its time stamp or incorporates the document into the catenate certificate.

A hashing function provides just such assurance, since at the time a document, such as an author's original work or a composite receipt catenation, is hashed there is created a representative "fingerprint" of its original content from which it is virtually impossible to recover that document. Therefore, the time-stamped document is not susceptible to revision by any adversary of the author. Nor is the author able to apply an issued time-stamp certificate to a revised form of the document, since any change in the original document content, even to the extent of a single word or a single bit of digital data, results in a different document that would hash to a completely different fingerprint value. Although a document cannot be recovered from its representative hash value, a purported original document can nonetheless be proven in the present time-stamping procedure by the fact that a receipt comprising a true copy of the original document representation will always hash to the original number or the same catenate value as is contained in the author's certificate, assuming use of the original hashing algorithm.

Any available deterministic function, e.g. a one-way hash function such as that described by Rivest ("The MD4 Message Digest Algorithm", Advances in Cryptology -- Crypto '90, Springer-Verlag, LNCS, to appear), incorporated herein by reference, may be used in the present procedure. In the practice of the invention, such a hashing operation is optionally employed by the author

to obtain the noted benefit of transmission security, although it might be effected by the TSA if the document were received in plaintext form. In whatever such manner the document content and incorporated time data are fixed against revision, there remains the further step, in order to promote the credibility of the system, of certifying to the members of an as yet unidentified universe that the receipt was in fact prepared by the TSA, rather than by the author, and that the time indication is correct, i.e., that it has not, for instance, been fraudulently stated by the TSA in collusion with the author.

To satisfy the former concern, the TSA uses a verifiable signature scheme, of a type such as the public key method earlier noted, to certify the time-stamp prior to its transmittal to the author. Confirmation of the signature at a later time, such as by decryption with the TSA's public key, proves to the author and to the universe at large that the certificate originated with the TSA. Proof of the veracity of the time-stamp itself, however, relies upon a following additional aspect of the invention.

In an alternative procedure, the TSA maintains a record of its sequential time-stamping transactions by adding each new receipt to its current catenation and applying its deterministic function, e.g. hashing, the composite to obtain a new catenation. This catenation, itself a value resulting from the hashing process, is included on the receipt or certificate returned to the author and serves to certify the indicated time stamp. Confirmation of the certificate at a later time involves rehashing the combination of the author's time receipt and the next previous catenate value in the TSA records. The resulting generation of the author's catenate certificate

value proves to the author and to the universe at large that the certificate originated with the TSA. This result also proves the veracity of the time-stamp itself, since all original elements of the original receipt must be repeated in order to again generate, by the hashing function, the original catenate certificate value.

One embodiment of the process, as generally depicted in FIG. 2, draws upon the relatively continuous flow of documents from the universe of authors through the facilities of the TSA. For each given processed document, D_k, the TSA generates a time-stamp receipt which includes, for example, a sequential receipt number, r_k, the identity of the author, A_k, by ID number, ID_k, or the like, the hash, H_k, of the document, and the current time, t_k. In addition, the TSA includes the receipt data of the immediately preceding processed document, D_{k-1}, of author, A_{k-1}, thereby bounding the time-stamp of document, D_k, in the "past" direction by the independently established earlier receipt time, t_{k-1}. Likewise, the receipt data of the next received document, D_{k+1}, are included to bound the time-stamp of document, D_k, in the "future" direction. The composite receipt, now containing the time data of the three, or more if desired, sequential time-stamp receipts, or identifying segments thereof, is then certified with the cryptographic TSA signature and transmitted to the author, A_k. In like manner, a certificate containing identifiable representations of D_k and D_{k+2} would be transmitted to author,A_{k+1}. Thus, each of the time-stamp certificates issued by the TSA is fixed in the continuum of time and none can be falsely prepared by the TSA, since a comparison of a number of relevant distributed certificates would reveal the discrepancy in their sequence. So effective is such a sequential fixing of a document in the time stream that the TSA signature could

be superfluous in actual practice.

A second embodiment of the invention, shown generally in FIG. 3, distributes the time-stamping task randomly among a broad universe, for example the multiplicity of authors utilizing the time-stamping process. A TSA could still be employed for administrative purposes or the requesting author could communicate directly with the selected time-stamping author/agents. In either event, the above-mentioned need for assurance that a time-stamp has not been applied to a document through collusion between the author and the stamping agency is met in the combination of the reasonable premise that at least some portion of the agency universe is incorruptible or would otherwise pose a threat of exposure to an author attempting falsification, and the fact that the time-stamping agencies for a given document are selected from the universe entirely at random. The resulting lack of a capability on the part of the author to select a prospective collusive agent of the author's own choosing substantially removes the feasibility of intentional time falsification.

The selection of the individual universe members who will act as the predetermined number of agents is accomplished by means of a pseudorandom generator of the type discussed by Impagliazzo, Levin, and Luby ("Pseudorandom Generation From One-Way Functions", Proc. 21st STOC, pp. 12-24, ACM, 1989) for which the initial seed is a deterministic function, such as a hash, of the document being time-stamped. Given as a seed input the document hash or other such function, the implemented pseudorandom generator will output a series of agency IDs. This agency selection is for all practical purposes unpredictable and random.

Once the agents are selected, the time-stamping proceeds as previously indicated with the exception that each agent individually adds the current time data to the representative document it receives, certifies the resulting separate time-stamped receipt with its own verifiable cryptographic signature, and transmits the certificate back to the author. This transmittal may be directly to the requesting author or by way of the administrative TSA where the receipts are combined with or without further certification by the TSA. The combination of signature scheme and a published directory of author IDs provides verification of the utilization of the agents that were in fact selected by the pseudorandom generator. This distributed agent embodiment of the invention presents some advantages over the receipt-linking procedure in that a certified time-stamp is provided more quickly and a given author's later proof of a document is less reliant upon the availability of the certificates of other authors.

In an additional embodiment shown in FIG. 4, the TSA generates a time-stamp receipt which includes, for example, a sequential receipt transaction number, r_k, the identity of the author, for example by ID number, ID_k, or the like, a digital representation, e.g. the hash, H_k, of the document, and the current time, t_k. The TSA then includes these receipt data, or any representative part thereof, with the catenate certificate value, C_{k-1}, of the immediately preceding processed document, D_{k-1}, of author, A_{k-1}, thereby bounding the time-stamp of document, D_k, by the independently established earlier receipt time, t_{k-1}.

The composite data string, $(r_k, ID_k, H_k, t_k, C_{k-1})$, is then hashed to a new catenate value, C_k, that is entered with transaction number, r_k, in the records of the TSA,

and is also transmitted to A_k, as the catenate certificate value, with the time-stamp receipt data. In like manner, a certificate value derived from the hashing of C_k with time stamp elements of the receipt for document, D_{k+1}, would be transmitted to author, A_{k+1}. Thus, each of the time-stamped catenate certificates issued by the TSA is fixed in the continuum of time and none can be falsely prepared by the TSA, since any attempt to regenerate a catenate certificate number from a hash with the next prior certificate would reveal the discrepancy.

In a more general application of the invention, as shown in FIG. 5, the representation, e.g., a hash, of a particular document is simply concatenated with the catenate certificate value of the next previous document and the deterministic function representation, again a hash, for example, of this composite is then generated and retained as the record catenate value for the particular document. Each subsequent document in the growing series is similarly processed to expand the record which itself would serve as a reliable certification of the position each such document occupies in the series, or more broadly viewed, in the continuum of time. This embodiment of the invention provides a reliable method by which an organization, for instance, could readily certify the sequence and continuity of its digital business documents and records.

Additional variations in the process of the invention might include the accumulation of documents, preferably in hashed or other representative form, generated within an author organization over a period of time, e.g. a day or more depending upon the extent of activity, with the collection being hashed to present a single convenient document for time-stamping and